

REMARKS

Claims 2-7 and 9-35 are currently pending in the application. Claims 2, 9, and 15 are independent claims. Claims 2-7 and 9-20 have been amended. New claims 21-35 have been added herein. Reconsideration of this application is respectfully requested.

Amendments to Specification

The specification has been amended at page 5 to remedy a typographical error that would be apparent to one of ordinary skill in the art, and at page 11 to remedy a printing error in the previous Amendment, wherein the mathematical symbol “θ” did not print properly. No new matter has been added.

Examiner’s Comments on Office Action Response Filed on November 15, 2005

The Examiner has suggested that the Applicant amend the claims to specify “selecting a vector profile,” because the Examiner contends that “the features upon which applicant relies (i.e., ‘a profile representing sports (P. 14, lines 22-25)’) are not recited in the rejected claim(s).” However, Applicant respectfully disagrees with the Examiner’s assertion.

Applicant wishes to clarify that the statement relating to sports in the previous Amendment was merely exemplary and was presented to reflect an example wherein, if a profile represents “sports”, a corresponding document score would be indicative of that content.

Applicant’s arguments did not rely on any specific type of profile, because the claims are not limited in this manner.

Art Rejections

The Office Action includes a rejection of claims 2, 3, 9, 10, 15, and 16 under 35 U.S.C. § 102(e) as allegedly being anticipated by U.S. Patent No. 6,189,002 issued to Roitblat *et al.* (hereinafter “Roitblat”). Claims 2, 3, 9, 10, 15, and 16 have been amended. The amendments to claims 2, 9 and 15 are for readability, to provide proper antecedent basis, and to correct a typographical error. These changes are not intended to be related to

patentability. As discussed below, it is respectfully submitted that claims 2, 3, 9, 10, 15, and 16 are not anticipated by Roitblat, and withdrawal of the rejection is respectfully requested.

Independent claim 2 recites a method of selecting documents from a data stream. The method comprises selecting a profile, analyzing a reference corpus of documents against the profile to determine at least one document score indicative of content of at least one document of the reference corpus relative to the profile, and scoring at least one data stream document from the data stream against said profile to provide a document score indicative of content in the data stream document relative to the profile. The method also comprises comparing the document score from the data stream document to the at least one document score from the analysis of the reference corpus to select the data stream document from the data stream. Claims 9 and 15 similarly recite, among other things, analyzing a reference corpus of documents against the profile to determine at least one document score indicative of content of at least one document of the reference corpus relative to the profile, and comparing the document score of the data stream (or data source) document to at least one document score from the analysis of the reference corpus to select the data stream (or data source) document.

In contrast, Roitblat does not disclose analyzing a reference corpus of documents against a profile to determine at least one document score indicative of content of at least one document of the reference corpus relative to the profile, nor does Roitblat disclose comparing a document score of a data stream (or data source) document to at least one document score from the analysis of the reference corpus to select the data stream (or data source) document as recited in the present independent claims. Rather, Roitblat discloses a method of selecting a document where a semantic profile generated from the user's query is compared to the semantic profiles of the new documents, such as world wide web pages. *See, e.g.,* Roitblat, col. 2, ll. 43-49. A neural network is trained to recognize the semantic profiles of terms using

the reference corpus texts, extract the context specific meaning of terms appearing in the corpus, and to extract vocabulary and semantic patterns. *See, e.g.*, Roitblat, col. 2, ll. 30-34, col. 3, ll. 38-40, and col. 3, ll. 52-55. Any new documents retrieved from, *e.g.*, the web, are submitted to the neural network for processing to produce their semantic profiles using the rules learned from the reference corpus texts. *See, e.g.*, Roitblat, col. 2, ll. 34-39. The user's query must similarly be transformed into a semantic profile. *See, e.g.*, Roitblat, col. 2, ll. 43-47. Thereafter, in order to select documents, Roitblat's method requires comparing the semantic profile generated from the user's query to the semantic profiles generated for the new documents. *See, e.g.*, Roitblat, col. 2, ll. 47-49.

The Office Action cites col. 3, ll. 20-65, col. 4, ll. 5-55, and col. 4, l. 55 – col. 5, l. 55 of Roitblat as allegedly disclosing analyzing a reference corpus of documents against a profile to determine at least one document score indicative of document content relative to the profile (the Office cites the latter section in particular as allegedly relating to document scoring). Applicant respectfully disagrees with the Office's characterization of Roitblat in this regard. Col. 4, l. 55- col. 5, l. 55 of Roitblat merely discusses aspects of training a neural network using a reference corpus (the discussion of which begins at col. 3, l. 20). Among the aspects discussed in the section cited by the Office include the construction of text vectors (col. 4, l. 54 – col. 5, l. 31) and the layer structure and learning approach of the neural network (col. 5, ll. 32-55). Nowhere in the cited sections, or elsewhere, does Roitblat disclose analyzing a reference corpus of documents against a profile to determine at least one document score indicative of content of at least one document of a reference corpus relative to the profile. Accordingly, since Roitblat does not disclose determining such a score, Roitblat cannot disclose comparing such a score to a document score of a data stream (or data source) document to select (or retrieve) the data stream (or data source) document as claimed in claims 2, 9 and 15. Should the Office disagree, the Office is respectfully requested to point

out precisely where Roitblat allegedly discloses determining at least one document score indicative of content of at least one document of a reference corpus relative to a profile.

The Office Action also cites col. 6, ll. 10-30 of Roitblat for allegedly disclosing comparing the document score of the data stream (or data source) document to at least one score from the reference corpus. However, this section of Roitblat merely discloses comparing text vectors derived directly from retrieved documents to text vectors generated by the neural network after submitting the retrieved document to it. *See* Roitblat, col. 6, ll. 12-17. The purpose for this comparison is for assessing the ability of the neural network to reproduce the document text vectors. *See* Roitblat, col. 6, ll. 10-12. There is no disclosure in the cited section of Roitblat of comparing the document score of the data stream (or data source) document to the document score from the analysis of the reference corpus to select (or retrieve) the data stream (or data source) document.

For at least the above-noted reasons, it is respectfully submitted that claims 2, 9 and 15 are not anticipated by Roitblat. Claims 3, 10 and 16 are allowable at least by virtue of dependency. Withdrawal of the rejection and allowance of claims 2, 3, 9, 10, 15, and 16 are respectfully requested.

The Office Action also includes a rejection of claims 4, 5, 11, 12, 17, and 18 under 35 U.S.C. § 103(a) as allegedly being obvious over Roitblat in view of U.S. Patent No. 6,327,574 (hereinafter “Kramer”). The Office Action cites Kramer for allegedly disclosing a method and system of comparing documents to topic profiles to determine closeness to a topic and the use of an exponential decay function scoring system (citing col. 1, l. 1 – col. 3, l. 60, and col. 30, l. 8-10). The Office further states that Roitblat discloses using alternative, well known statistical techniques for the normalization of vectors and scoring systems (citing col. 7, ll. 25-27; col. 8, ll. 11-16 and 48-55).

As an initial matter, it is respectfully pointed out that the Office's reliance on Kramer does not make up for the deficiencies of Roitblat as discussed above in connection with claims 2, 9 and 15. Thus, claims 4, 5, 11, 12, 17 and 18 are allowable at least by virtue of dependency. Withdrawal of the rejection and allowance of claims 4, 5, 11, 12, 17 and 18 are respectfully requested.

In addition, Applicant respectfully disagrees with the Office's characterization of Roitblat and Kramer in connection with the above-noted sections of these patents. Kramer is directed to selecting and electronically delivering documents to a consumer based on the consumer's profile, which includes hierarchical attribute vectors which encode attributes of a consumer (Abstract). Col. 30, ll. 8-11 of Kramer contain no disclosure of "exponential decay scoring systems", contrary to the Office's suggestion, and merely disclose that updating attribute vectors (of the consumer profile) may be done with Bayesian updating methods or other techniques such as exponential decay, wavelets, and Gaussian combination.

Further, Applicant sees no disclosure of "usage of alternative, well known statistical techniques for the normalization of vectors and scoring systems" in the cited sections of Roitblat (col. 7, ll. 25-27; col. 8, ll. 11-16 and 48-55), contrary to the Office's suggestion. Col. 7, ll. 25-27 of Roitblat merely states, "Two embodiments are described below in detail. Additional embodiments using alternative neural networks are also briefly mentioned." Col. 8, ll. 11-16 merely states, "The identical transform [referring to a mathematical transform in the prior paragraph] can be accomplished using well known statistical techniques. The network and statistical techniques differ only in the details of the algorithm by which the N principal components are computed, the results are identical." Applicants note that the preceding paragraph of Roitblat regarding the "transform" (col. 7, l. 53 – col. 8, l. 11) contains no disclosure regarding document scoring. Further, col. 8, ll. 48-55 of Roitblat merely states, "There are other techniques that can be used in place of principal components

to project the high dimensional text vectors onto lower dimensional representations. These techniques are known in the statistical literature as matrix decomposition techniques. They differ somewhat in the constraints they place on the projected dimensions, but most of them would yield good results for information retrieval.” Again, the cited section contains no disclosure regarding document scoring. Claims 4, 5, 11, 12, 17 and 18 are additionally allowable for at least the above-noted reasons.

The Office Action also includes a rejection of claims 6, 7, 13, 14, 19, and 20 under 35 U.S.C. § 103(a) as allegedly being obvious over Roitblat in view of the Nature Article entitled “Diameter of the Web” (hereinafter “the Nature Article”). The Office Action cites to the Nature Article for allegedly disclosing the use of power-law functions in document analysis. The Office further states that Roitblat discloses using alternative, well known statistical techniques for the normalization of vectors and scoring systems (citing col. 7, ll. 25-27; col. 8, ll. 11-16 and 48-55 as above), and cites Roitblat at col. 5, lines 8-14, as allegedly disclosing usage of Zipf’s law as being related to power law methods.

As an initial matter, it is respectfully pointed out that the Office’s reliance on the Nature Article does not make up for the deficiencies of Roitblat as discussed above in connection with claims 2, 9 and 15. Thus, claims 6, 7, 13, 14, 19, and 20 are allowable at least by virtue of dependency. Withdrawal of the rejection and allowance of claims 6, 7, 13, 14, 19, and 20 are respectfully requested.

In addition, Applicant respectfully disagrees with the Office’s characterization of Roitblat and the Nature Article in connection with the above-noted sections of these documents. Applicant’s disagreement with the Office’s characterization of Roitblat at col. 7, ll. 25-27; col. 8, ll. 11-16 and 48-55 is discussed above. In addition, the Office’s reliance upon Roitblat at col. 5, lines 8-14 regarding the reference with Zipf’s law is misplaced, because this section merely discusses the size of the corpus and states that the frequency of

occurrence of a work is a decreasing logarithmic function of the rank of the work. The section has nothing to do with using a power law or logarithmic function in document scoring. In addition, the Nature Article relates to links to Internet documents and states that the probabilities that an Internet document has k outgoing and incoming links follows a power law over several orders of magnitude (page 1, 3rd paragraph). The mention of “power law” in this context has nothing to do with using a power law in document scoring. Accordingly, the Office’s reliance on this document is also misplaced. Claims 6, 7, 13, 14, 19, and 20 are additionally allowable for at these additional reasons.

New Claims 21-35

New claims 21-35 have been added herein to round out the scope of protection sought. Support for these claims may be found, for example, at least at pages 4, 5, 6 and 11 of the present application and in the originally filed claims. It is respectfully submitted that these claims are allowable at least for reasons similar to those set forth above, and further because these claims recite additional distinguishing subject matter not disclosed in the applied references. Allowance of claims 21-35 is respectfully requested.

CONCLUSION

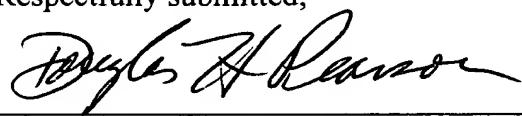
In light of the above, withdrawal of the rejections and allowance of the application are respectfully requested. Should any issues remain in connection herewith, the Examiner is invited to telephone the undersigned to discuss the same.

No fee is believed due in connection with this response. In the event that a fee is required, please charge any such fees to Jones Day Deposit Account No. 50-3013.

Respectfully submitted,

Date:

June 2, 2006


Douglas H. Pearson (Reg. No. 47,851)
JONES DAY
51 Louisiana Ave., NW
Washington, DC 20001-2113
(202) 879-3939